

4L-11

知的インタビューシステム I²S に基づく データベース構築支援システムの開発

沼田 薫 川口 敦生 溝口 理一郎 山口 高平 角所 収
(大阪大学産業科学研究所)

1. はじめに

本稿では、データベースの論理設計を支援する知的インタビューシステム I²S を応用したデータベース構築支援システム I²DC (Intelligent Interview based Database Construction Support System) について述べる。I²DC は I²S を包含しており、既存の DBMS の知的なフロントエンドとして機能する。

本システムは、データベース構築作業全般にわたって、これをマニュアルレス化することを目標としている。従来人間が担当してきた論理設計作業までも知的インタビューシステム I²S が担当するため、この目標を十分に達成できると考えている。

2. システムの構成とデータベース構築手順

システムの構成を図1に示す。システムの動作は、supervisor によって管理される。論理設計以降の作業は、ルール形式で表現された知識を用いて、プロダクションシステムが行う。

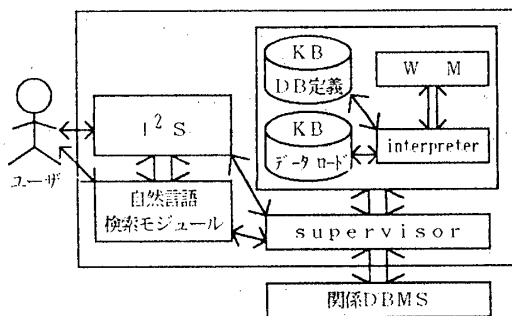


図1 システムの構成

本システムによるデータベース構築手順を以下に示す。

- 1) I²S がユーザーにインタビューを行ない、論理設計を行う。
- 2) 設計結果より対象DBMSの DDL (Database Definition Language) を用いてデータベースを定義、生成する。
- 3) 論理設計結果、データベースの定義より初期データロード用のプログラムを生成する。同時にデータロード用のファイルのフォーマットが決定される。

- 4) ユーザーに3)のファイルフォーマットを提示し、初期データファイルを作成してもらう。
- 5) 4)のデータファイルを3)のプログラムを用いてロードする。データベースの第1版が完成する。
- 6) I²DC が提供する自然言語による検索機能を利用して、ユーザーはデータベースをアクセスする。(インプリメントは検討中)

以下、個々の処理について順に述べる。ただし処理1については別稿⁽¹⁾を参照されたい。例として用いる I²S の論理設計結果を図2に示す。対象DBMSとしては、Prolog でインプリメントしたDBMSシミュレータを用いる。このシミュレータは、SQL型のDML (Data Manipulation Language) を処理できる。

```

ent(suppliers, [], [sno, sname, status, scity]).
ent(parts, [], [pno, pname, color, weight, pcity]).
ent(department, [], [dept, dcity, manager]).
key(suppliers, [sno]).
key(parts, [pno]).
key(department, [dept]).
rel(shipment, [suppliers, parts], [sh_qty]).
rel(stock, [department, parts], [st_qty]).
hie(suppliers, sno, '1:1').
hie(suppliers, sname, '1:1').
hie(suppliers, status, '1:1').
hie(suppliers, scity, '1:1').
hie(parts, pno, '1:1').
hie(parts, pname, '1:1').
hie(parts, color, '1:1').
hie(parts, weight, '1:1').
hie(parts, pcity, '1:1').
hie(department, dept, '1:1').
hie(department, dcity, '1:1').
hie(department, manager, '1:1').
type(suppliers, sno, char).
type(suppliers, sname, char).
type(suppliers, status, int).
type(suppliers, scity, char).
type(parts, pno, char).
type(parts, pname, char).
type(parts, color, char).
type(parts, weight, real).
type(parts, pcity, char).
type(department, dept, char).
type(department, dcity, char).
type(department, manager, char).
type(shipment, sh_qty, int).
type(stock, st_qty, int).
    
```

図2 論理設計結果の例

3. データベースの生成

データベースの定義、生成は、I²S の設計結果を対象

DBMSのDDLに基づいて変換し、DBMSに引き渡すことによって行う。本システムでは、変換規則をルール形式で表現し、前向き推論を行うプロダクションシステムを用いて変換を行う。図3に変換ルールの例を示す。図4に図2を変換した例を示す。

```

if
  e_table(Table),
  hie(Table,Attribute,'1:1'),
  type(Table,Attribute,Type),
  attribute_list(Attribute_list),
  type_list(Type_list)
then
  remove_wm(2),
  modify_wm(4,attribute_list([Attribute|Attribute_list])),
  modify_wm(5,type_list([Type|Type_list])).
if
  ent(Table,_,_)
then
  remove_wm(1),
  make_wm(e_table(Table)),
  make_wm(attribute_list({})),
  make_wm(type_list({})).

```

図3 設計結果→データベース定義変換ルールの例

```

database_definition
  table suppliers
    attributes
      sno: char_not_null
      sname: char
      status: int
      scity: char
    key sno
  table parts
    attributes
      pno: char_not_null
      pname: char
      color: char
      weight: real
      pcity: char
    key pno
  table department
    attributes
      dept: char_not_null
      dcity: char
      manager: char
    key dept
  table shipment
    attributes
      sno: char_not_null
      pno: char_not_null
      sh_qty: int
    key sno,pno
  table stock
    attributes
      dept: char_not_null
      pno: char_not_null
      st_qty: int
    key dept,pno
end_database_definition.

```

図4 図2の例をデータベース定義に変換した例

4. 初期データのロード

論理設計結果から、まずデータファイルのフォーマットを決定し仕様書を出力する。次にデータロード用プログラムを生成する。どちらも3と同様、前向き推論を行うプロダクションシステムを用いて行う。図5に図2の例において出力される仕様書を示す。

5. 自然言語による検索

自然言語による検索機能は、現在検討中である。問題点としては

Input data from datafile for table "suppliers",
"parts","department","shipment" and "stock".
Datafile format is as follows.

```

pno (char)
pname (char)
color (char)
weight (real)
pcity (char)
< Now, data for 1 parts is ended. >
< You can choose from the followings. >
< 1. Put more data by the same format >
< 2. Put a NEW_LINE code to show the end of block >
sno (char)
sname (char)
status (int)
scity (char)
pno (char)
sh_qty (int)
:
< same format for each shipment >
:
NEW_LINE < the delimiter for shipment >
< Now, data for 1 suppliers is ended. >
< You can choose from the followings. >
< 1. Put more data by the same format >
< 2. Put a NEW_LINE code to show the end of block >
dept (char)
dcity (char)
manager (char)
pno (char)
st_qty (int)
:
< same format for each stock >
:
NEW_LINE < the delimiter for stock >
< Now, data for 1 department is ended. >
< You can choose from the followings. >
< 1. Put more data by the same format >
< 2. Put a NEW_LINE code to show the end of block >

```

図5 データフォーマット仕様書の例

1. 解析時の曖昧さ
2. 類義語の取扱い
3. 不合理な検索結果が生じた際の応答^[2]

といった事柄が挙げられる。これらの問題は、ユーザが持っているデータベースシステムに対するイメージと実際のシステムとのミスマッチに起因している。本システムは知的インタビューシステムI²Sを包含しているので、これを活用することによってミスマッチの解消に当たることを考えている。

6. 終りに

本稿では、知的インタビューシステムI²Sに基づくデータベース構築支援システムI²DCの、処理の流れおよび自然言語による検索機能の基本的考え方を述べた。

システムは、日本データゼネラル社のMV/8000II上のMV-Prologを用いてインプリメントされている。今後、同社の関係型DBMS、DG/SQLを対象DBMSとするシステムも開発する予定である。

参考文献:

- [1] 川口他, 「知的インタビューシステムI²Sとその質問戦略」, 情報処理学会第33回全国大会, 4L-10, 1986.
- [2] Kaplan, S. J., "Cooperative Responses From a Portable Natural Language Database Query System," Brady, M. and Berwick, R. C. (eds.), Computational Models of Discourse, The MIT Press, 1983.